Jyot Buch

Champaign, Illinois | jyotb2@illinois.edu | +1 217 819 9416

Education

University of Illinois, Urbana-Champaign, MSc in Statistics — Aug 2024 – May 2026

- GPA: 3.8/4.0
- **Coursework:** Causal Inferences, Deep Learning for Computer Vision, Unsupervised Learning

Vishwakarma University, Pune, BTech in Computer Engineering — Aug 2019 – May 2023

- GPA: 8.7/10.0
- **Coursework:** Database Management Systems, Big Data Analytics, Data Mining and Predictive Modeling

Technical Skills

- **Programming Languages:** Python, R, SQL, C++, Java, CSS, HTML, Javascript, HiveQl
- **Tools & Technologies:** Power BI, Streamlit, Tableau, Apache Hadoop, OpenCV, AWS Bedrock
- **Machine Learning:** PyTorch, Transformers, LLMs, Unsupervised and Supervised learning approaches

Work Experience

Data Science Intern @ *Amazon* — May 2025 – Aug 2025

- Built FinLens, a **document-agnostic entity extraction** pipeline using **AWS Bedrock, S3, and Lambda** that consolidated cross-team prompt engineering efforts, reducing development time while maintaining >95% accuracy on diverse financial documents through **advanced prompt engineering and confidence gating**
- Enhanced Minerva **RAG system** for tax analysts by integrating **multimodal capabilities** with **dense captioning models, implementing citation support** for source traceability, and improving **BM25 Recall@10 by 8%** through systematic chunking strategy experimentation
- Designed **novel intent-aware completeness recall metric** for summarization evaluation that maps **user instructions to document facts**, enabling data-driven model selection and providing actionable feedback for **improving recall**.

Automation Developer @ *Accelirate Softech* — Dec 2023 – Jul 2024

- Built an **intelligent invoice processing** workflow using **UiPath Agents**, **leveraging computer vision models and language models** to automatically extract vendor information, line items, and payment terms from diverse document formats.
- Developed **SQL procedures** to cross-reference extracted contract terms against predefined **compliance standards** stored in reference databases, **automatically flagging deviations** and generating exception reports.

Data Science Intern @ *Eaton* — Aug 2022 – Dec 2022

- Developed the Drawing-to-Data (D2D) tool using **YOLOv5 and Tesseract OCR** to **automate information extraction** from 250,000+ engineering drawings, achieving over **91% detection accuracy** and **93% OCR precision**
- Developed an automated compliance mapping system using **Python, MySQL, and VBA** to analyze client requirements against organizational historical data, generating comprehensive **compliance matrices** and streamlining requirement analysis workflows.
- Recognized with IP (**Intellectual Property**) for the D2D tool.

Assistantships

Teaching and Research Assistant @ *University of Illinois, Urbana-Champaign*

- Served as Teaching Assistant for **BADM 554: Enterprise Database Management**, providing hands-on coaching to students on database management systems, SQL optimization, and enterprise data architecture principles.
- Built a **Graph RAG pipeline** linking **PostgreSQL and Neo4j** to create a live **student–assessment–error knowledge graph**; integrated **LLMs** for personalized, **concept-tagged feedback** with **direct lecture references** and performance tracking.
- Implemented **LocalStack for AWS service emulation** to train students, and **introduced the "Right-of-AI" learning stack** (Claude for reasoning, Cursor for coding, Marimo for notebooks) to support onboarding and development workflows.

Projects

Ball Action Recognition in Soccer Videos using CNN-Transformer Hybrid on SoccerNet Dataset

- Built a **CNN-Transformer hybrid model** on the SoccerNet dataset for ball action detection (pass, shot, save), implementing **2.5D frame processing** with spatial-temporal cues and **frame stacking with temporal attention** mechanisms.
- Optimized model performance through **cosine LR scheduling**, label smoothing, and multi-scale temporal **feature fusion**, achieving **0.713 mAP@0.75** score comparable to established benchmarks with enhanced action recognition accuracy.